

COMBINED WAVELET DOMAIN AND MOTION COMPENSATED FILTERING COMPLIANT WITH VIDEO CODECS

*Ljubomir Jovanov¹, Aleksandra Pižurica¹, Vladimir Zlokolica¹, Stefaan Schulte²,
Etienne Kerre² and Wilfried Philips¹*

¹ Telecommunications and Information Processing
Ghent University
St-Pietersnieuwstraat 41
B-9000 Ghent, Belgium
ljj@telin.ugent.be

² Fuzziness and Uncertainty Modelling Group
Ghent University
Krijgslaan 281 - S9
B-9000 Ghent, Belgium
Stefaan.Schulte@ugent.be

ABSTRACT

In this paper, we introduce the idea of using motion estimation resources from a video codec for video denoising. This is not straightforward because the motion estimators aimed for video compression and coding, tolerate errors in the estimated motion field and hence are not directly applicable to video denoising. To solve this problem, we propose a novel motion field filtering step that refines the accuracy of the motion estimates to a degree that is required for denoising.

We illustrate the use of the proposed motion estimation method within a wavelet-based video denoising scheme. The resulting video denoising method, is of low-complexity and compares favorably to the latest video denoising methods.

Index Terms— video, denoising, motion estimation, compression, coding

1. INTRODUCTION

A number of wavelet-based video denoising algorithms have been proposed recently, including [1], [2], [3] and [4]. In a closely related video processing chain, wavelet-based video codecs prove their advantages, especially in terms of resolution scalability.

In many applications such as video surveillance or tele-medicine, it should be desirable to integrate video denoising and video coding, as closely related parts of the same video processing chain. A natural step towards this integration is re-using the resources, such as motion estimation, from a video codec for the denoiser. A number of efficient motion estimators have been already developed, such as [5], which are suitable for real-time hardware implementation [6]. However, the use of these motion estimators has not been considered so far in the denoising literature.

The main reason is that motion estimator aimed for video compression and coding tolerate errors in the estimated motion field. Often, motion field produced by video codec does not follow the actual moving object which is one of the conditions required to perform successful motion-compensated denoising.

In this paper, we propose a simple and efficient filtering step which allows integration of a motion estimator from a video codec into a denoiser. In particular, we observe that inaccuracies of multiresolution motion estimators, such as [5] consist mainly in false motion vectors in background image areas without actual motion. The proposed, novel motion field filtering step refines the accuracy of the motion field making it usable for denoiser. We show that in

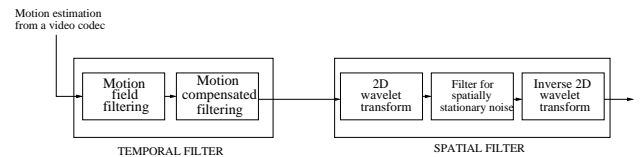


Fig. 1. Tested video denoising scheme.

combination with a wavelet based spatial filter scheme the proposed video denoising scheme competes with the best and most recent multiresolution video denoisers, such as [3]. While promoting the idea of integrating the motion estimator from a video codec into a denoiser, we do not restrict ourselves to a particular codec or denoiser design. The essence is that we can use the output of the same motion estimator as an input for the coding scheme and with the proposed additional filtering step, as an input to the denoiser.

We also develop a novel wavelet-based video denoising scheme that fits in the proposed video processing strategy. The results demonstrate that the proposed filter outperforms recent related methods such as [3] and [4], while being compatible with video coding motion estimation algorithms. The paper is organized as follows. Section 2 describes the proposed spatio-temporal filtering scheme. The results are presented and discussed in Section 3. Conclusions are given in Section 4.

We adopt a separable and sequential spatio-temporal filtering scheme, related to [4], which combines wavelet-domain spatial filtering and pixel-domain temporal filtering. The main difference with respect to [4] is a novel, motion compensated temporal filter, as opposed to simple pixel based motion detection in [4]. Secondly we use a different filtering order, where the temporal filtering precedes spatial one. This is because we observed that such a scheme yields less spatial blur as compared to the filtering order from [4]. One difficulty with our temporal-spatial filter ordering is that the remaining noise after the temporal filtering is in general spatially non-stationary, which requires adaptivity of the spatial filter to local noise statistics. Different solutions for this problem are possible, including the extension of the spatial filter from [7]. Here we propose a related solution for non-stationary spatial filter, starting from a fuzzy-logic version of this filter [8], which offers a good compromise between denoising quality and complexity, being attractive for hardware implementation. We extend this filter to make it adaptive to spatially

non-stationary noise.

2. THE PROPOSED VIDEO DENOISER

We assume the following noise model:

$$d_{i,j}^k = f_{i,j}^k + n_{i,j}^k, \quad (1)$$

where (i, j) denote the spatial position, k -the frame number, $f_{i,j}^k$ the noise-free component and $n_{i,j}^k$ the noise contribution. Noise is assumed to be white Gaussian of zero mean and of known standard deviation σ . In case σ is not known, we estimate it as described in [3]. The proposed video denoising scheme is illustrated in Fig. 1. An input into this scheme is the motion field produced by a motion estimator from a video codec. In particular we assume the multiresolution spatio temporal motion estimation algorithm (MRST) described in [5]. This algorithm uses a multigrid image representation and starts with a full search on the lowest resolution, and uses higher resolution levels and spatially and temporally neighboring blocks to improve the existing motion vectors in an adaptive manner. This motion estimation algorithm offers similar results as a full search algorithm but with multiple speed-up.

We have noticed that the main problem with this, and other motion estimation methods aimed for video compression is that they produce motion vectors in the parts of the frame where no motion exists (for example in the background which contains steep illumination gradients). This is because the corresponding motion estimation algorithms search for the best matching pixel value in the previous frame no matter if it does not belong to the same object. Such motion field cannot be used for video denoising without further refinement, since spurious motion vectors would introduce degradations of stationary parts of the scene.

2.1. The proposed motion filtering step

In order to eliminate spurious motion vectors, from the parts of image sequence where no motion exist we propose a novel motion field filtering step. The idea is to compare the difference between the corresponding blocks with the average difference, and decide whether motion exists or not. In the first step we calculate mean absolute difference between pixels in the corresponding blocks in neighboring frames, denoted with $D_{i,j}^k$:

$$D_{i,j}^k = \frac{1}{N^2} \sum_{m=1}^{N_x} \sum_{n=1}^{N_y} |d_{m,n}^{k,i,j} - d_{m,n}^{k-1,i,j}|, \quad (2)$$

where i, j are spatial coordinates of a block, m, n coordinates of a pixel inside block used for motion estimation and N is a block size.

We define the threshold used for motion detection in k -th frame as follows:

$$THR = \gamma \frac{1}{N_{bh} N_{bv}} \sum_{i=1}^{N_{bh}} \sum_{j=1}^{N_{bv}} D_{i,j}^k, \quad (3)$$

where γ is a scalar and N_{bh}, N_{bv} are number of blocks along horizontal and vertical axis. We found experimentally that value $\gamma = 0.55$ yields best results for the most sequences.

In this filtering step, we make decision whether motion exists in each block simply by comparing the absolute block difference with the previously calculated threshold. If the $D_{i,j}^k < THR$, both motion vector components are set to zero. Otherwise, motion vector keeps its original value.

2.2. The proposed motion compensated temporal filtering

Let $m_{i,j}$ denote the motion detection variable at position i, j in the frame k . Value $m_{i,j}^k = 1$ indicates “motion” and $m_{i,j}^k = 0$ “no motion” at the corresponding spatio-temporal position. In the original SEQWT method from [4], the recursive temporal filtering is simply switched off at positions where motion is detected ($m_{i,j}^k = 1$). Using our motion estimator, we shall filter at the positions where motion exists as well, and this along the motion trajectory. We wish to take into account however, that the motion estimation is not perfect and hence we apply different coefficients of the filter in moving and in non-moving areas. The proposed motion compensated filter is:

$$\begin{aligned} \hat{f}_{i,j}^k = & (1 - m_{i,j}^k) \left(\alpha \left(1 + \frac{MAD}{255} \right) d_{i,j}^k + \left(1 - \frac{MAD}{255} \right) (1 - \alpha) d_{i,j}^{k-1} \right) \\ & + m_{i,j}^k \left(\beta \left(1 + \frac{MAD}{255} \right) d_{i,j}^k + \left(1 - \frac{MAD}{255} \right) (1 - \beta) d_{i-p,j-q}^{k-1} \right), \end{aligned} \quad (4)$$

where (p, q) are the coordinates of the motion vector and $MAD = |d_{i,j}^k - d_{i-p,j-q}^{k-1}|$. Optimal values for filter coefficients $\alpha=0.55$ and $\beta=0.9$ are chosen experimentally to yield best denoising results for all test sequences.

2.3. Spatial filtering for non-stationary noise

Like in [4], we combine the temporal filter with a wavelet domain spatial filter. As opposed to [4], our spatial filter follows the temporal one, and hence it has to deal with spatially non-stationary noise. Aiming at low complexity and hardware-friendly solution, we start from the fuzzy filter of [8], which is a fuzzy-logic version of the spatially adaptive ProbShrink from [9]. The filter applies to each wavelet coefficient a shrinkage factor, which is a function of two measurements: the coefficient magnitude and a local spatial activity indicator (LSAI). These functions are realized as piece-wise linear membership functions which depend on standard deviation of noise σ . For details see [8]. We make this method adaptive to spatially non-stationary noise by estimating σ locally. We use 32x32 overlapping windows and shift these in steps of 8 pixels along each direction.

3. RESULTS

We first analyze the performance of the proposed motion estimation algorithm and next the performance of the complete denoising algorithm.

3.1. Motion field improvement

The performance of the algorithm is evaluated by comparing mean square error of the motion compensated frame, obtained using the estimated motion field with and without the proposed motion field filtering step. Fig. 3 shows the motion field for 15-th frame of ‘chair’ sequence before and after filtering.

Mean square error of the motion field is defined as

$$MSE = \frac{1}{N_x N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} (f_{i,j}^k - f_{i-p,j-q}^{k-1})^2, \quad (5)$$

N_x, N_y are the sizes of image frame along x and y axis, i, j coordinates of pixel inside image frame and p, q are the x and y components of the motion vector. The resulting mean square errors given

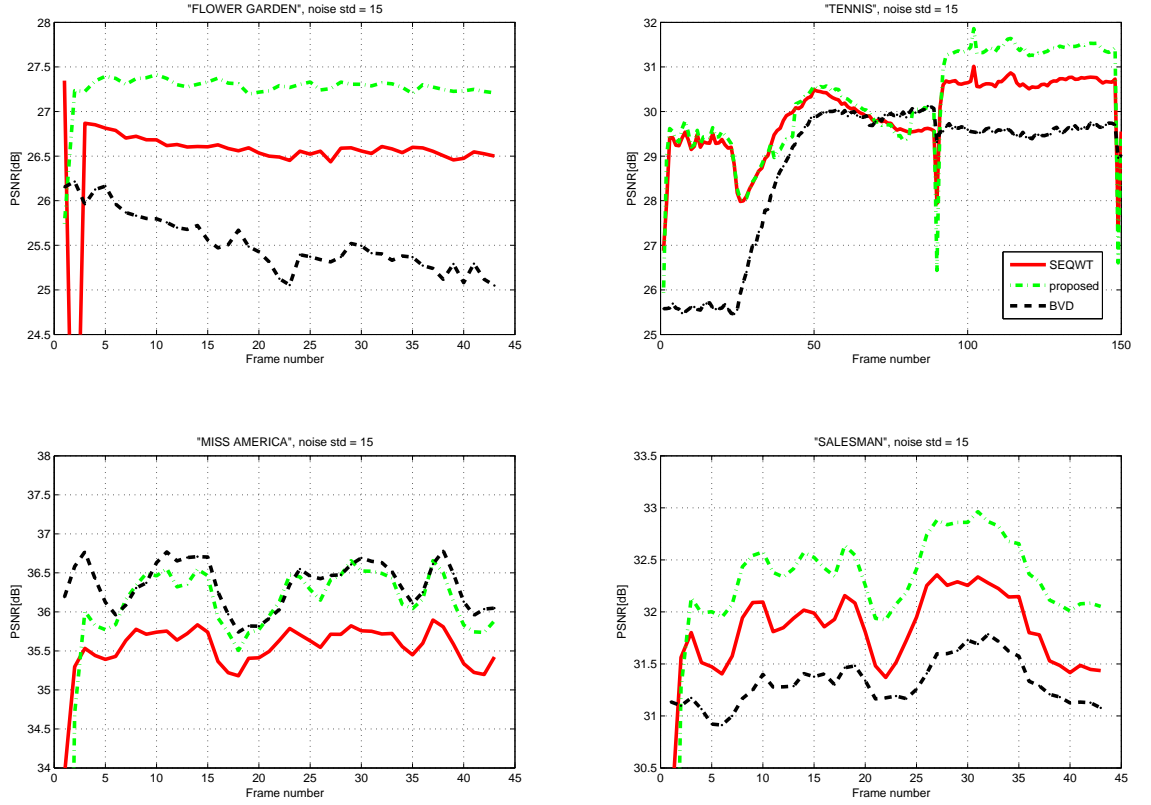


Fig. 2. Quantitative performance comparison of the proposed method and SEQWT and BVD algorithms for the following sequences: (a) flower, $\sigma=15$ (b) tennis, $\sigma=15$ (c) miss America, $\sigma=15$ (d) salesman, $\sigma=15$

Table 1. Mean square error of the motion field

Sequence	Without MV filtering	With MV filtering
"salesman"	0.188	0.118
"chair"	0.077	0.047
"miss America"	0.062	0.036
"tennis"	0.017	0.017
"flower garden"	0.49	0.345
"bus"	0.286	0.202

in Table. 1 demonstrate that the proposed filtering step improves the accuracy of the estimated motion field.

3.2. Denoising results

In this section we present the denoising performance of the proposed method. The wavelet function used here is *symlet* with eight vanishing moments. All the results were obtained using a non-decimated wavelet transform with 4 decomposition levels.

Four test sequences were used, 'miss America', 'salesman', 'tennis' and 'flower', with added white Gaussian noise of standard deviations $\sigma=10, 15$ and 20 . Resulting PSNR of the proposed method

is compared with the methods presented in [3] and [4], denoted as BVD and SEQWT respectively. which is illustrated in Fig. 2. The proposed denoising method shows better results for all test sequences. The improvements in PSNR are from 0.5 - 0.8 dB compared with [4], depending on the sequence. These improvements are mainly because the new method includes a motion compensated filtering. A significant factor in performance improvement is also the use of a spatially adaptive local filtering which adapts itself to non-stationary noise. The visual results shown in Fig. 4, demonstrate that the proposed method also outperforms the reference methods in terms of visual quality, especially in the sequences which contain more dynamics. Differences between the presented methods are even more perceivable after viewing and comparing denoised sequences.

4. CONCLUSION

The main novelty of this paper is the idea of reusing motion estimation resources from video codec for video denoising. The benefit of the proposed idea is twofold. Firstly we improve the performance of the SEQWT denoiser [4] by using a motion estimator that is suitable for real-time processing. Secondly we make a first step towards integrating wavelet domain video coder and denoiser by making them

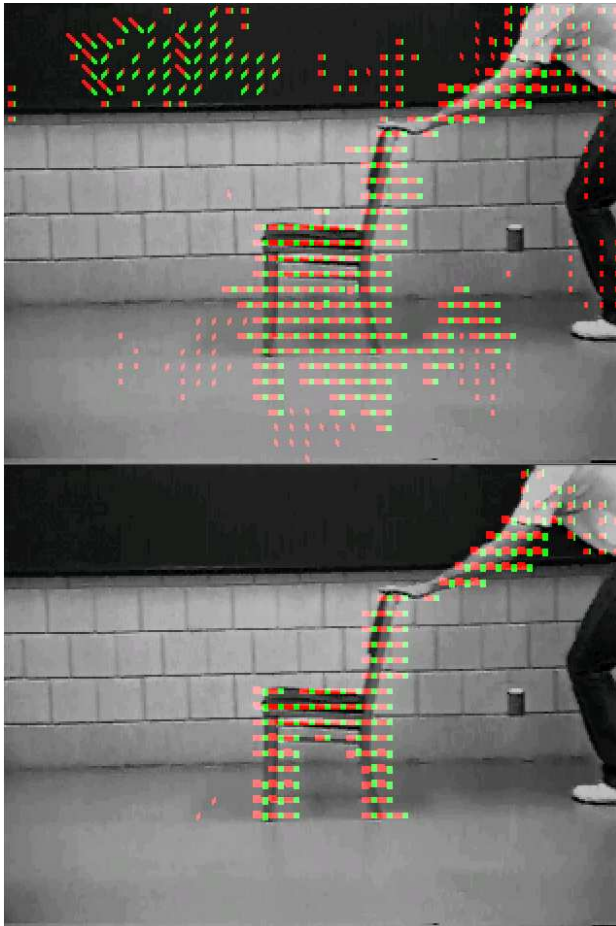


Fig. 3. Motion field for 15-th frame of 'chair' sequence, before and after filtering.

sharing the common resources such as motion estimation. This is not straightforward since the motion estimators aimed for video compression tolerate errors in the estimated motion field and hence are not applicable to video denoising. In order to achieve this, we introduce novel motion filtering step that improves significantly the mean error squared performance of the motion estimator. This configuration can be easily implemented as a part of hardware video codec, since both components have low complexity.

5. REFERENCES

- [1] V. Zlokolica, A. Pižurica, and W. Philips, "Wavelet-domain video denoising based on reliability measures," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 7, no. 3, pp. 477–488, 2006.
- [2] W.I. Selesnick and K.Y. Li, "Video denoising using 2d and 3d dualtree complex wavelet transforms," in *SPIE Wavelet Applications in Signal and Image Processing*, San Diego, USA, 2003, vol. 5207.
- [3] E. J. Balster, Y. F. Zheng, and R. L. Ewing, "Combined spatial and temporal domain wavelet shrinkage algorithm for video

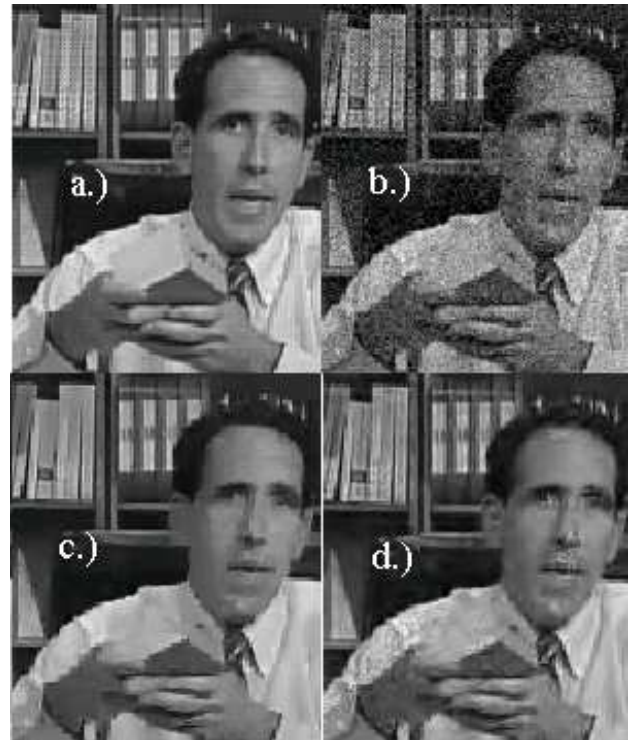


Fig. 4. Fragment of the 7-th frame of the 'Salesman': a.) original, b.) noisy sequence with $\sigma = 15$, c.) BVD algorithm [1], d.) proposed method

denoising," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 2, pp. 220–230, Feb. 2006.

- [4] A. Pižurica, V. Zlokolica, and W. Philips, "Noise reduction in video sequences using wavelet-domain and temporal filtering," in *SPIE Conference "Wavelet Applications in Industrial Processing"*, Providence, Rhode Island, USA, 2003.
- [5] J. Chalidabhongse and C.-C. Jay Kuo, "Fast motion vector estimation using multiresolution-spatio-temporal correlations," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 16, no. 8, pp. 993–1007, 2006.
- [6] J. Hsieh and T. Meng, "Low-power parallel video compression architecture for a single-chip digital cmos camer," *Journal of VLSI Signal Processing Systems*, vol. 21, no. 11, pp. 195–207, July 1999.
- [7] B. Goossens, A. Pižurica, and W. Philips, "Wavelet domain image denoising for non-stationary noise and signal dependent noise," in *IEEE International Conf. On Image Processing (ICIP)*, Atlanta, GA, USA, 2006.
- [8] S. Schulte, B. Huysmans, Pižurica, E. Kerre, and W. Philips, "A new fuzzy-based wavelet shrinkage image denoising technique," in *Advanced Concepts for Intelligent Vision Systems (Acivs 2006)*, Antwerp, Belgium, 2006.
- [9] A. Pižurica and W. Philips, "Estimating probability of presence of a signal of interest in multiresolution single- and multiband image denoising," *IEEE Trans. on Image Processing*, vol. 15, no. 3, pp. 654–665, 2006.